AD A030654

13

REPEATED MEASUREMENTS ON AUTOREGRESSIVE PROCESSES

BY

T. W. ANDERSON

TECHNICAL REPORT NO. 25
SEPTEMBER 1976

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

D D C
OCT 13 1976
C

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

REPEATED MEASUREMENTS ON AUTOREGRESSIVE PROCESSES

By

T. W. Anderson

Technical Report No. 25
September 1976

PREPARED UNDER CONTRACT N00014-75-C-0442
(NR-042-034)
OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A

Repeated Measurements on Autoregressive Processes

by

T. W. Anderson
Stanford University

1. Introduction.

Time series analysis is developed mainly for the study of a single sequence of observations. In this paper methods of statistical inference are developed for a set of several sequences of observations from the same autoregressive vector-valued process. Having several time series available from the same model permits investigation of such questions as whether the process is homogeneous, that is, whether the autoregression coefficients are constant in time; if the coefficients vary in time they can, nevertheless, be estimated from several series even though there may be no pattern to the variation.

The distributions of estimates and test criteria in time series analysis are usually asymptotic as the length of the series increases. In the present case of repeated measurements, asymptotic theory can alternatively be based on the number of observed series increasing. In cases where both large-sample theories might be used they are compatible.

There are many applications of these procedures of importance. In a panel survey (in which several respondents are interviewed at more than one point in time) the responses may be quantitative, such as answers to the question how many hours did you spend last month reading the newspapers. In economic surveys the questions are likely to produce numerical answers: how many hours did you work last week and how much money did you spend on groceries last month. Analysis of such data are sometimes called cross-section studies by econometricians. A psychologist may obtain a test score on several individuals at several dates; a physician may read blood pressures of his patients on several days.

The purpose of the present paper is to develop statistical methods for problems of autoregressive processes in the case of repeated measurements which are analogs of procedures for corresponding problems of Markov chains treated by Anderson and Goodman (1957). Anderson (1976) reviewed some of these procedures for Markov chains and suggested analogs for autoregressive processes. This present paper develops those and other procedures in greater detail and more generally and provides justification for some assertions in the earlier paper.

2. The Model.

If $\underset{\sim}{y}_t$ is a p-component vector, a Markov (first-order) vector process with mean $\mathcal{E}\underset{\sim}{y}_t = \underset{\sim}{0}$ is defined by

$$\underset{\sim}{y}_t = \underset{\sim}{B}(t)\underset{\sim}{y}_{t-1} + \underset{\sim}{u}_t \ , \tag{2.1}$$

where $\underset{\sim}{B}(t)$ is a $p \times p$ matrix and $\{\underset{\sim}{u}_t\}$ is a sequence of independent (unobservable) random vectors with expected values $\mathcal{E}\underset{\sim}{u}_t = \underset{\sim}{0}$, covariance matrices $\mathcal{E}\underset{\sim}{u}_t\underset{\sim}{u}_t' = \underset{\sim}{\Sigma}_t$, and $\underset{\sim}{u}_t$ independent of $\underset{\sim}{y}_{t-1}, \underset{\sim}{y}_{t-2}, \ldots$. Let the covariance matrix of $\underset{\sim}{y}_t$ be $\mathcal{E}\underset{\sim}{y}_t\underset{\sim}{y}_t' = \underset{\sim}{F}_t$. Then from (2.1) and the independence of $\underset{\sim}{y}_{t-1}$ and $\underset{\sim}{u}_t$ we deduce

$$\underset{\sim}{F}_t = \underset{\sim}{B}(t)\, \underset{\sim}{F}_{t-1}\, \underset{\sim}{B}(t)' + \underset{\sim}{\Sigma}_t \ . \tag{2.2}$$

If the observations are made for $t = 1, \ldots, T$, the model may be specified by the marginal distribution of $\underset{\sim}{y}_1$ and the distributions of $\underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_T$. In particular, if $\underset{\sim}{y}_1$ and the $\underset{\sim}{u}_t$'s are normal, the model for the observation period is specified by $\underset{\sim}{F}_1, \underset{\sim}{B}(2), \ldots, \underset{\sim}{B}(T), \underset{\sim}{\Sigma}_2, \ldots, \underset{\sim}{\Sigma}_T$.

When the autoregression matrices are homogeneous, that is, $\underset{\sim}{B}(t) = \underset{\sim}{B}$, and the $\underset{\sim}{u}_t$'s are identically distributed with mean $\underset{\sim}{0}$ and covariance matrice $\underset{\sim}{\Sigma}$, the process is stationary if the characteristic roots of $\underset{\sim}{B}$ are less than 1 in absolute value and the process is defined for $t = \ldots, -1, 0, 1, \ldots$ or if $\underset{\sim}{y}_1$ is assigned the stationary marginal distribution. In this case the covariance matrix of

$\underset{\sim}{y}_t$ is

$$\underset{\sim}{F} = \sum_{s=0}^{\infty} \underset{\sim}{B}^s \underset{\sim}{\Sigma} \underset{\sim}{B}'^s , \tag{2.3}$$

and the covariance matrix of $\underset{\sim}{y}_t$ and $\underset{\sim}{y}_s$ is

$$\mathcal{E}\underset{\sim}{y}_t \underset{\sim}{y}_s' = \underset{\sim}{B}^{t-s} \underset{\sim}{F} , \qquad s \leq t . \tag{2.4}$$

[Note that $\underset{\sim}{F} = \underset{\sim}{F}_t = \underset{\sim}{F}_{t-1}$ satisfies (2.2) for $\underset{\sim}{B}(t) = \underset{\sim}{B}$ and $\underset{\sim}{\Sigma}_t = \underset{\sim}{\Sigma}$, $t = 2, \ldots, T$.]

A higher-order process may be an appropriate model. The r-th order inhomogeneous process is

$$\underset{\sim}{y}_t = \sum_{j=1}^{r} \underset{\sim}{B}_j(t) \underset{\sim}{y}_{t-j} + \underset{\sim}{u}_t , \tag{2.5}$$

and the homogeneous process is

$$\underset{\sim}{y}_t = \sum_{j=1}^{r} \underset{\sim}{B}_j \underset{\sim}{y}_{t-j} + \underset{\sim}{u}_t . \tag{2.6}$$

In the latter case the roots of

$$\left| -\lambda^r \underset{\sim}{I} + \sum_{j=1}^{r} \lambda^{r-j} \underset{\sim}{B}_j \right| = 0 \tag{2.7}$$

shall be less than 1 in absolute value.

A higher-order autoregressive process may be written as a first-order process by suitable redefinition. For example, a second-order autoregressive vector process with homogeneous autoregressive coefficients defined by

$$\underset{\sim}{y}_t = \underset{\sim}{B}_1 \underset{\sim}{y}_{t-1} + \underset{\sim}{B}_2 \underset{\sim}{y}_{t-2} + \underset{\sim}{u}_t \tag{2.8}$$

can be written as a first-order process by

$$\tilde{\underset{\sim}{y}}_t = \tilde{\underset{\sim}{B}} \tilde{\underset{\sim}{y}}_{t-1} + \tilde{\underset{\sim}{u}}_t \ , \tag{2.9}$$

where

$$\tilde{\underset{\sim}{y}}_t = \begin{pmatrix} \underset{\sim}{y}_t \\ \underset{\sim}{y}_{t-1} \end{pmatrix}, \quad \tilde{\underset{\sim}{u}}_t = \begin{pmatrix} \underset{\sim}{u}_t \\ \underset{\sim}{0} \end{pmatrix}, \tag{2.10}$$

$$\tilde{\underset{\sim}{B}} = \begin{pmatrix} \underset{\sim}{B}_1 & \underset{\sim}{B}_2 \\ \underset{\sim}{I} & \underset{\sim}{0} \end{pmatrix} \ . \tag{2.11}$$

The characteristic roots of $\tilde{\underset{\sim}{B}}$ are the roots of

$$|-\lambda^2 \underset{\sim}{I} + \lambda \underset{\sim}{B}_1 + \underset{\sim}{B}_2| = 0 \ . \tag{2.12}$$

For a stationary process these roots are to be less than 1 in absolute value.

The autoregressive processes appropriate to several subpopulations (strata) may be different. In the inhomogeneous first-order case the matrices in the h-th process may be $\underset{\sim}{B}^{(h)}(t)$, $\underset{\sim}{\Sigma}_t^{(h)}$, $h = 1, \ldots, s$; in the homogeneous case they may be $\underset{\sim}{B}^{(h)}$, $\underset{\sim}{\Sigma}^{(h)}$, $h = 1, \ldots, s$. If influencing variables are continuous, they may be taken account of by adding them to the regression to yield the model (in the first-order homogeneous case)

$$\underset{\sim}{y}_t = \underset{\sim}{B}\, \underset{\sim}{y}_{t-1} + \underset{\sim}{\gamma}'\, \underset{\sim}{z}_t + \underset{\sim}{u}_t \,, \tag{2.13}$$

where $\underset{\sim}{z}_t$ is a vector of such variables and $\underset{\sim}{\gamma}$ is a vector of parameters. In particular, when $\underset{\sim}{z}_t \equiv 1$ and $\underset{\sim}{\gamma}$ is a scalar, the process $\{\underset{\sim}{y}_t\}$ may have a mean different from $\underset{\sim}{0}$.

The autoregressive process is constructed from (conditional) multivariate regressions. In (2.1), for example, the vector $\underset{\sim}{y}_{t-1}$ constitutes the "independent variables" and the vector $\underset{\sim}{y}_t$ constitutes the "dependent variables" in ordinary regression. To a large extent the statistical methods for autoregressive models are regression or least squares procedures.

3. Estimation of Autoregressive Coefficients.

Let $\underset{\sim}{y}_{t\alpha}$ be the p-component vector of measurements of the α-th individual at the t-th time point, $\alpha = 1, \ldots, N$, $t = 1, \ldots, T$. The model is a first-order autoregressive model (2.1) with $\underset{\sim}{u}_t$ having the normal distribution $N(\underset{\sim}{0}, \underset{\sim}{\Sigma}_t)$ and $\underset{\sim}{y}_1$ having the normal distribution $N(\underset{\sim}{0}, \underset{\sim}{F}_1)$. The probability density of the sequences $\underset{\sim}{y}_{1\alpha}, \ldots, \underset{\sim}{y}_{T\alpha}$, $\alpha = 1, \ldots, N$, is

$$(3.1) \quad \prod_{\alpha=1}^{N} \frac{1}{(2\pi)^{\frac{1}{2}Tp} |\underset{\sim}{F}_1|^{\frac{1}{2}} \prod_{t=2}^{T} |\underset{\sim}{\Sigma}_t|^{\frac{1}{2}}}$$

$$\times \exp\left\{-\frac{1}{2}\left[\underset{\sim}{y}_{1\alpha}' \underset{\sim}{F}_1^{-1} \underset{\sim}{y}_{1\alpha} + \sum_{t=2}^{T} (\underset{\sim}{y}_{t\alpha} - \underset{\sim}{B}(t)\underset{\sim}{y}_{t-1,\alpha})' \underset{\sim}{\Sigma}_t^{-1} (\underset{\sim}{y}_t - \underset{\sim}{B}(t)\underset{\sim}{y}_{t-1,\alpha})\right]\right\}$$

$$= \frac{1}{(2\pi)^{\frac{1}{2}NTp} |\underset{\sim}{F}_1|^{\frac{1}{2}N} \prod_{t=2}^{T} |\underset{\sim}{\Sigma}_t|^{\frac{1}{2}N}} \exp\left\{-\frac{1}{2} \sum_{\alpha=1}^{N} \left[\underset{\sim}{y}_{1\alpha}' \underset{\sim}{F}_1^{-1} \underset{\sim}{y}_{1\alpha}\right.\right.$$

$$\left.\left. + \sum_{t=2}^{T} (\underset{\sim}{y}_{t\alpha} - \underset{\sim}{B}(t)\, \underset{\sim}{y}_{t-1,\alpha})' \underset{\sim}{\Sigma}_t^{-1} (\underset{\sim}{y}_{t\alpha} - \underset{\sim}{B}(t)\, \underset{\sim}{y}_{t-1,\alpha})\right]\right\}$$

$$= \frac{1}{(2\pi)^{\frac{1}{2}NTp} |\underset{\sim}{F}_1|^{\frac{1}{2}N} \prod_{t=2}^{T} |\underset{\sim}{\Sigma}_t|^{\frac{1}{2}N}} \exp\left\{-\frac{1}{2}\left[\operatorname{tr} \underset{\sim}{F}_1^{-1} \sum_{\alpha=1}^{N} \underset{\sim}{y}_{1\alpha} \underset{\sim}{y}_{1\alpha}'\right.\right.$$

$$+ \sum_{t=2}^{T} \operatorname{tr} \underset{\sim}{\Sigma}_t^{-1} \left(\sum_{\alpha=1}^{N} \underset{\sim}{y}_{t\alpha} \underset{\sim}{y}_{t\alpha}' - 2\underset{\sim}{B}(t) \sum_{\alpha=1}^{N} \underset{\sim}{y}_{t-1,\alpha} \underset{\sim}{y}_{t\alpha}'\right.$$

$$\left.\left.\left. + \underset{\sim}{B}(t) \sum_{\alpha=1}^{N} \underset{\sim}{y}_{t-1,\alpha} \underset{\sim}{y}_{t-1,\alpha}' \underset{\sim}{B}(t)'\right)\right]\right\} .$$

Then $1/N$ times the logarithm of the likelihood function is

$$(3.2)\quad \frac{1}{N}\log L = -\frac{1}{2}T\log 2\pi - \frac{1}{2}\log|\underset{\sim}{F}_1| - \frac{1}{2}\sum_{t=2}^{T}\log|\underset{\sim}{\Sigma}_t| - \frac{1}{2}\Big\{\operatorname{tr}\underset{\sim}{F}_1^{-1}\underset{\sim}{C}_1(0)$$

$$+\sum_{t=2}^{T}\operatorname{tr}\underset{\sim}{\Sigma}_t^{-1}[\underset{\sim}{C}_t(0) - 2\,\underset{\sim}{C}_t(1)\,\underset{\sim}{B}'(t) + \underset{\sim}{B}(t)\,\underset{\sim}{C}_{t-1}(0)\,\underset{\sim}{B}'(t)]\Big\} ,$$

where

$$(3.3)\qquad \underset{\sim}{C}_t(j) = \frac{1}{N}\sum_{\alpha=1}^{N}\underset{\sim}{y}_{t\alpha}\,\underset{\sim}{y}'_{t-j,\alpha} .$$

Then a sufficient set of statistics for $\underset{\sim}{F}_1, \underset{\sim}{B}(2), \ldots, \underset{\sim}{B}(T), \underset{\sim}{\Sigma}_2, \ldots, \underset{\sim}{\Sigma}_T$ is $\underset{\sim}{C}_t(0)$, $t = 1, \ldots, T$, and $\underset{\sim}{C}_t(1)$, $t = 2, \ldots, T$.

The derivatives of $(1/N)\log L$ with respect to the elements of $\underset{\sim}{B}(t)$ form the matrix

$$(3.4)\qquad \frac{1}{N}\frac{\partial \log L}{\partial \underset{\sim}{B}(t)} = -\underset{\sim}{\Sigma}_t^{-1}[\underset{\sim}{C}_t(1) - \underset{\sim}{B}(t)\,\underset{\sim}{C}_{t-1}(0)] , \qquad t = 2, \ldots, T .$$

The maximum likelihood estimates of $\underset{\sim}{B}(2), \ldots, \underset{\sim}{B}(T)$ are

$$(3.5)\qquad \hat{\underset{\sim}{B}}(t) = \underset{\sim}{C}_t(1)\,\underset{\sim}{C}_{t-1}^{-1}(0) , \qquad t = 2, \ldots, T .$$

Then $1/N$ times the logarithm of the concentrated likelihood function is

$$(3.6) \qquad -\frac{1}{2} Tp \log 2\pi - \frac{1}{2} \log |\underset{\sim}{F}_1| - \frac{1}{2} \sum_{t=2}^{T} \log |\underset{\sim}{\Sigma}_t|$$

$$- \frac{1}{2}\left\{ \operatorname{tr} \underset{\sim}{F}_1^{-1} \underset{\sim}{C}_1(0) + \sum_{t=2}^{T} \operatorname{tr} \underset{\sim}{\Sigma}_t^{-1}\left[\underset{\sim}{C}_t(0) - \hat{\underset{\sim}{B}}(t)\, \underset{\sim}{C}_{t-1}(0)\, \hat{\underset{\sim}{B}}(t)'\right]\right\} .$$

Then [by Lemma 3.2.2 of Anderson (1958)] the maximum likelihood estimates of $\underset{\sim}{F}_1, \underset{\sim}{\Sigma}_2, \ldots, \underset{\sim}{\Sigma}_T$ are

$$(3.7) \qquad \hat{\underset{\sim}{F}}_1 = \underset{\sim}{C}_1(0) ,$$

$$(3.8) \qquad \hat{\underset{\sim}{\Sigma}}_t = \underset{\sim}{C}_t(0) - \hat{\underset{\sim}{B}}(t)\, \underset{\sim}{C}_{t-1}(0)\, \hat{\underset{\sim}{B}}(t)'$$

$$= \underset{\sim}{C}_t(0) - \underset{\sim}{C}_t(1)\, \underset{\sim}{C}_{t-1}(0)^{-1}\, \underset{\sim}{C}_t(1)'$$

$$= \frac{1}{N} \sum_{\alpha=1}^{N} \left(\underset{\sim}{y}_{t\alpha} - \hat{\underset{\sim}{B}}(t)\underset{\sim}{y}_{t-1,\alpha}\right)\left(\underset{\sim}{y}_{t\alpha} - \hat{\underset{\sim}{B}}(t)\underset{\sim}{y}_{t-1,\alpha}\right)', \quad t = 2, \ldots, T .$$

The components of $\underset{\sim}{B}(t)$ are least squares estimates and $\hat{\underset{\sim}{\Sigma}}_t$ are composed of sums of products of deviations from the fitted (auto) regressions. [See Anderson (1958), Chapter 8, and Anderson (1971), Chapter 5, for example.]

The assumption that the autoregression matrices are homogeneous and the disturbances identically distributed leads to considerable simplification. Then $1/N$ times the logarithm of the likelihood function is

$$(3.9)\qquad \frac{1}{N}\log L = -\frac{1}{2}\,Tp\log 2\pi - \frac{1}{2}\log|\underset{\sim}{F}_1| - \frac{1}{2}(T-1)\log|\underset{\sim}{\Sigma}|$$

$$-\frac{1}{2}\Bigg\{\operatorname{tr}\underset{\sim}{F}_1^{-1}\underset{\sim}{C}_1(0) + \operatorname{tr}\underset{\sim}{\Sigma}^{-1}\Bigg[\sum_{t=2}^{T}\underset{\sim}{C}_t(0)$$

$$-2\sum_{t=2}^{T}\underset{\sim}{C}_t(1)\,\underset{\sim}{B}' + \underset{\sim}{B}\sum_{t=2}^{T}\underset{\sim}{C}_{t-1}(0)\,\underset{\sim}{B}'\Bigg]\Bigg\}$$

$$= -\frac{1}{2}\,Tp\log 2\pi - \frac{1}{2}\log|\underset{\sim}{F}_1| - \frac{1}{2}(T-1)\log|\underset{\sim}{\Sigma}|$$

$$-\frac{1}{2}\Bigg\{\operatorname{tr}\underset{\sim}{F}_1^{-1}\underset{\sim}{C}_1(0) + (T-1)\operatorname{tr}\underset{\sim}{\Sigma}^{-1}\Bigg[\underset{\sim}{C}(0) - 2\,\underset{\sim}{C}(1)\,\underset{\sim}{B}'$$

$$+\underset{\sim}{B}\,\underset{\sim}{C}^{*}(0)\,\underset{\sim}{B}'\Bigg]\Bigg\},$$

where

$$(3.10)\qquad \underset{\sim}{C}(j) = \frac{1}{T-1}\sum_{t=2}^{T}\underset{\sim}{C}_t(j)\ ,$$

$$(3.11)\qquad \underset{\sim}{C}^{*}(0) = \frac{1}{T-1}\sum_{t=2}^{T}\underset{\sim}{C}_{t-1}(0)$$

$$= \underset{\sim}{C}(0) + \frac{1}{T-1}\left[\underset{\sim}{C}_1(0) - \underset{\sim}{C}_T(0)\right]\ .$$

Then a sufficient set of statistics for $\underset{\sim}{F}_1$, $\underset{\sim}{B}$, and $\underset{\sim}{\Sigma}$ is $\underset{\sim}{C}_1(0)$, $\underset{\sim}{C}(0)$, $\underset{\sim}{C}^{*}(0)$ and $\underset{\sim}{C}(1)$. The maximum likelihood estimate of $\underset{\sim}{F}_1$ is (3.7), and the maximum likelihood estimates of $\underset{\sim}{B}$ and $\underset{\sim}{\Sigma}$ are

$$\hat{\underset{\sim}{B}} = \underset{\sim}{C}(1)\ \underset{\sim}{C}^*(0)^{-1}\ , \tag{3.12}$$

$$\begin{aligned}\hat{\underset{\sim}{\Sigma}} &= \underset{\sim}{C}(0) - \hat{\underset{\sim}{B}}\ \underset{\sim}{C}^*(0)\ \hat{\underset{\sim}{B}}' \\ &= \underset{\sim}{C}(0) - \underset{\sim}{C}(1)\ \underset{\sim}{C}^*(0)^{-1}\ \underset{\sim}{C}(1)' \\ &= \frac{1}{N(T-1)} \sum_{t=2}^{T} \sum_{\alpha=1}^{N} (\underset{\sim}{y}_{t\alpha} - \hat{\underset{\sim}{B}}\ \underset{\sim}{y}_{t-1,\alpha})(\underset{\sim}{y}_{t\alpha} - \hat{\underset{\sim}{B}}\ \underset{\sim}{y}_{t-1,\alpha})'\ .\end{aligned} \tag{3.13}$$

An alternative model is to consider $\underset{\sim}{y}_{1\alpha}$, $\alpha = 1, \ldots, N$, as nonstochastic or fixed and treat $\underset{\sim}{y}_{t\alpha}$, $t = 2, \ldots, T$, $\alpha = 1, \ldots, N$, conditionally. Then the maximum likelihood estimates of $\underset{\sim}{B}(t)$ and $\underset{\sim}{\Sigma}_t$, $t = 2, \ldots, T$, are (3.5) and (3.8) and of $\underset{\sim}{B}$ and $\underset{\sim}{\Sigma}$ are (3.12) and (3.13), as the case may be.

When $\underset{\sim}{y}_{1\alpha}$ is considered to have the marginal normal distribution determined by the stationary process, the covariance matrix $\underset{\sim}{F}_1$ is a function of $\underset{\sim}{B}$ and $\underset{\sim}{\Sigma}$ as given by (2.3). Then the maximum likelihood estimates are much more complicated. In the simplest case of $p = 1$ the estimation equations involve a cubic [Sec. 6.11.1 of Anderson (1971)]; if $p > 1$ or the order is greater than 1, polynomial equations of higher degree are involved. [As $T \to \infty$, (3.12) and (3.13) are asymptotically equivalent to the maximum likelihood estimates, but not as $N \to \infty$.]

Estimation in higher-order processes is similar. For example, the second-order case may be treated in the form (2.8) or (2.9). Suppose

$\underset{\sim}{y}_{1\alpha}, \ldots, \underset{\sim}{y}_{r\alpha}$, $\alpha = 1, \ldots, N$, are considered as nonstochastic or fixed. Then in the r-th order process the maximum likelihood estimates of $\underset{\sim}{B}_j(t)$, $j = 1, \ldots, r$, $t = r+1, \ldots, T$, are the solutions to

$$(3.14) \qquad \left[\hat{\underset{\sim}{B}}_1(t) \ \ldots \ \hat{B}_r(t)\right] = \left[\underset{\sim}{C}_t(1) \ \ldots \ C_t(r)\right] \begin{bmatrix} \underset{\sim}{C}_t(1,1) & \ldots & \underset{\sim}{C}_t(1,r) \\ \vdots & & \vdots \\ \underset{\sim}{C}_t(r,1) & \ldots & \underset{\sim}{C}_t(r,r) \end{bmatrix}^{-1},$$

where

$$(3.15) \qquad \underset{\sim}{C}_t(i,j) = \frac{1}{N} \sum_{\alpha=1}^{N} \underset{\sim}{y}_{t-i,\alpha} \, \underset{\sim}{y}'_{t-j,\alpha} \ .$$

The covariance matrix $\underset{\sim}{\Sigma}_t$ is estimated by

$$(3.16) \qquad \hat{\underset{\sim}{\Sigma}}_t = \frac{1}{N} \sum_{\alpha=1}^{N} \left[\underset{\sim}{y}_{t\alpha} - \sum_{i=1}^{r} \hat{\underset{\sim}{B}}_t(i) \, \underset{\sim}{y}_{t-i,\alpha} \right] \left[\underset{\sim}{y}_{t\alpha} - \sum_{j=1}^{r} \hat{\underset{\sim}{B}}_t(j) \, \underset{\sim}{y}_{t-j,\alpha} \right]'$$

$$= \underset{\sim}{C}_t(0) - \sum_{i,j=1}^{r} \hat{\underset{\sim}{B}}_t(i) \, \underset{\sim}{C}_t(i,j) \, \hat{\underset{\sim}{B}}(j)' \ .$$

In the homogeneous case the estimates are

$$(3.17) \qquad \left[\hat{\underset{\sim}{B}}_1 \ \ldots \ \hat{\underset{\sim}{B}}_r\right] = \left[\underset{\sim}{C}(1|r) \ \ldots \ \underset{\sim}{C}(r|r)\right] \begin{bmatrix} \underset{\sim}{C}(1,1|r) & \ldots & \underset{\sim}{C}(1,r|r) \\ \vdots & & \vdots \\ \underset{\sim}{C}(r,1|r) & \ldots & \underset{\sim}{C}(r,r|r) \end{bmatrix}^{-1} ,$$

where

$$(3.18) \qquad \underset{\sim}{C}(j|r) = \frac{1}{T-r} \sum_{t=r+1}^{T} \underset{\sim}{C}_t(j,j)$$

$$= \frac{1}{T-r} \sum_{t=r+1-j}^{T-j} \underset{\sim}{C}_t(0) ,$$

$$(3.19) \qquad \underset{\sim}{C}(i,j|r) = \frac{1}{T-r} \sum_{t=r+1}^{T} \underset{\sim}{C}_t(i,j) .$$

The exact distributions of these estimates are, in general, too complicated to be useful. Hence, we turn to asymptotic theory. For fixed T by the law of large numbers

$$(3.20) \qquad \operatorname*{plim}_{N\to\infty} \underset{\sim}{C}_t(0) = \mathcal{E}\underset{\sim}{y}_t \underset{\sim}{y}_t' = \underset{\sim}{F}_t , \qquad t = 1, \ldots, T ,$$

$$(3.21) \qquad \operatorname*{plim}_{N\to\infty} \underset{\sim}{C}_t(1) = \mathcal{E}\underset{\sim}{y}_t \underset{\sim}{y}_{t-1}' = \underset{\sim}{B}_t \underset{\sim}{F}_t \qquad t = 2, \ldots, T .$$

Hence, as $N \to \infty$, $\underset{\sim}{C}_t(0)$ is a consistent estimate of $\underset{\sim}{F}_t$, $t = 1, \ldots, T$; $\hat{\underset{\sim}{B}}(t)$ is a consistent estimate of $\underset{\sim}{B}(t)$, $t = 2, \ldots, T$, and $\hat{\underset{\sim}{\Sigma}}_t$ is a

consistent estimate of $\underset{\sim}{\Sigma}_t$, $t = 2, \ldots, T$. It follows from (3.20) and (3.21) that in the homogeneous case

$$\text{(3.22)} \qquad \operatorname*{plim}_{N\to\infty} \underset{\sim}{C}(0) = \frac{1}{T-1} \sum_{t=2}^{T} \underset{\sim}{F}_t ,$$

which is $\underset{\sim}{F}$ in the stationary case [$\underset{\sim}{F}_1 = \underset{\sim}{F}$ given by (2.3)],

$$\text{(3.23)} \qquad \operatorname*{plim}_{N\to\infty} \underset{\sim}{C}(1) = \frac{1}{T-1} \underset{\sim}{B} \sum_{t=2}^{T} \underset{\sim}{F}_{t-1} ,$$

which is $\underset{\sim}{B}\,\underset{\sim}{F}$ in the stationary case, and

$$\text{(3.24)} \qquad \operatorname*{plim}_{N\to\infty} \underset{\sim}{C}^*(0) = \frac{1}{T-1} \sum_{t=2}^{T} \underset{\sim}{F}_{t-1} ,$$

which is $\underset{\sim}{F}$ in the stationary case. It follows that in the homogeneous case, as $N \to \infty$, $\hat{\underset{\sim}{B}}$ is a consistent estimate of $\underset{\sim}{B}$ and $\hat{\underset{\sim}{\Sigma}}$ is a consistent estimate of $\underset{\sim}{\Sigma}$.

For fixed T as $N \to \infty$ the elements of $\sqrt{N}[\underset{\sim}{C}_t(0) - \mathcal{E}\underset{\sim}{C}_t(0)]$, $t = 1, \ldots, T$, and of $\sqrt{N}[\underset{\sim}{C}_t(1) - \mathcal{E}\underset{\sim}{C}_t(1)]$ have a joint limiting normal distribution with means 0 if it is assumed that $\underset{\sim}{y}_1$ is random and the set $(\underset{\sim}{y}_{1\alpha}, \underset{\sim}{u}_{2\alpha}, \ldots, \underset{\sim}{u}_{T\alpha})$ are identically distributed (with respect to $\alpha = 1, \ldots, N$) . It follows that $\sqrt{N}[\hat{\underset{\sim}{F}}_1 - \underset{\sim}{F}_1]$, $\sqrt{N}[\hat{\underset{\sim}{B}}(t) - \underset{\sim}{B}(t)]$, and $\sqrt{N}[\hat{\underset{\sim}{\Sigma}}_t - \underset{\sim}{\Sigma}_t]$ has a joint limiting normal distribution with means 0. It remains to find the covariances. The limiting distribution of

$$(3.25) \qquad \sqrt{N}[\hat{\underset{\sim}{B}}(t) - \underset{\sim}{B}(t)] = \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{N} \underset{\sim}{u}_{t\alpha} \, \underset{\sim}{y}'_{t-1,\alpha} \, \underset{\sim}{C}_{t-1}(0)^{-1}$$

is the limiting distribution of

$$(3.26) \qquad \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{N} \underset{\sim}{u}_{t\alpha} \, \underset{\sim}{y}'_{t-1,\alpha} \, \underset{\sim}{F}_{t-1}^{-1}$$

since $\operatorname{plim}_{N \to \infty} \underset{\sim}{C}_{t-1}(0) = \underset{\sim}{F}_{t-1}$. The covariance matrix of the i-th and j-th rows of (3.26) is

$$(3.27) \qquad \frac{1}{N} \mathcal{E} \sum_{\alpha,\gamma=1}^{N} u_{it\alpha} \, u_{jt\gamma} \, \underset{\sim}{F}_{t-1}^{-1} \, \underset{\sim}{y}_{t-1,\alpha} \, \underset{\sim}{y}'_{t-1,\gamma} \, \underset{\sim}{F}_{t-1}^{-1}$$

$$= \frac{1}{N} \sigma_{ij}^{(t)} \, \underset{\sim}{F}_{t-1}^{-1} \mathcal{E} \sum_{\alpha=1}^{N} \underset{\sim}{y}_{t-1,\alpha} \, \underset{\sim}{y}'_{t-1,\alpha} \, \underset{\sim}{F}_{t-1}^{-1}$$

$$= \sigma_{ij}^{(t)} \, \underset{\sim}{F}_{t-1}^{-1} \quad ,$$

where $\underset{\sim}{\Sigma}_t = [\sigma_{ij}^{(t)}]$. The covariances of the elements of $\sqrt{N}[\hat{\underset{\sim}{B}}(t) - \underset{\sim}{B}(t)]$ in the limiting distribution constitute the Kronecker product

$$(3.28) \qquad \underset{\sim}{\Sigma} \otimes \underset{\sim}{F}_{t-1}^{-1} = \begin{pmatrix} \sigma_{11} \, \underset{\sim}{F}_{t-1}^{-1} & \sigma_{12} \, \underset{\sim}{F}_{t-1}^{-1} & \cdots & \sigma_{1p} \, \underset{\sim}{F}_{t-1}^{-1} \\ \sigma_{21} \, \underset{\sim}{F}_{t-1}^{-1} & \sigma_{22} \, \underset{\sim}{F}_{t-1}^{-1} & \cdots & \sigma_{2p} \, \underset{\sim}{F}_{t-1}^{-1} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} \, \underset{\sim}{F}_{t-1}^{-1} & \sigma_{p2} \, \underset{\sim}{F}_{t-1}^{-1} & \cdots & \sigma_{pp} \, \underset{\sim}{F}_{t-1}^{-1} \end{pmatrix} .$$

The elements of (3.26) for different values of t are uncorrelated; the covariance matrix of the i-th row with t and the j-th row with t replaced by s is

$$(3.29) \qquad \frac{1}{N} \mathcal{E} \sum_{\alpha,\gamma=1}^{N} u_{it\alpha} u_{js\gamma} \underset{\sim}{F}_{t-1}^{-1} \underset{\sim}{y}_{t-1,\alpha} \underset{\sim}{y}'_{s-1,\alpha} \underset{\sim}{F}_{s-1}^{-1} = \underset{\sim}{0} , \quad t \neq s ;$$

if $t > s$, then $\underset{\sim}{y}_{t-1,\alpha} \underset{\sim}{y}'_{s-1,\alpha}$ consists of quadratic forms in components of $\underset{\sim}{u}_{r\alpha}$, $r = 1, \ldots, t-1$, and $\underset{\sim}{y}_{1\alpha}$ and the expected value of such a term multiplied by $u_{it\alpha} u_{js\gamma}$ is 0.

When T is fixed and $N \to \infty$

$$(3.30) \qquad \sqrt{N}(\hat{\underset{\sim}{B}} - \underset{\sim}{B}) = \frac{1}{\sqrt{N}(T-1)} \sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{u}_{t\alpha} \underset{\sim}{y}'_{t-1,\alpha} \underset{\sim}{C}^{*}(0)^{-1}$$

has the same limiting distribution as

$$(3.31) \qquad \frac{1}{\sqrt{N}(T-1)} \sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{u}_{t\alpha} \underset{\sim}{y}_{t-1,\alpha} \underset{\sim}{F}^{*-1} ,$$

where

$$(3.32) \qquad \underset{\sim}{F}^{*} = \frac{1}{T-1} \sum_{t=2}^{T} \underset{\sim}{F}_{t-1} .$$

The covariance matrix of the i-th and j-th rows of (3.31) is $[1/(T-1)]\sigma_{ij} \underset{\sim}{F}^{*-1}$. Thus the limiting distribution of $\sqrt{N}(\hat{\underset{\sim}{B}} - \underset{\sim}{B})$ is normal with mean $\underset{\sim}{0}$ and covariance matrix $\underset{\sim}{\Sigma} \otimes (\Sigma_{t=2}^{T} \underset{\sim}{F}_{t-1})^{-1}$. If the process is stationary $F^{*} = \underset{\sim}{F}$ and $\sqrt{N(T-1)}(\hat{\underset{\sim}{B}} - \underset{\sim}{B})$ has the

limiting normal distribution with mean $\underset{\sim}{0}$ and covariance matrix $\underset{\sim}{\Sigma} \otimes \underset{\sim}{F}^{-1}$.

Limiting distributions when N is fixed and $T \to \infty$ are of interest only in the homogeneous case. For $N = 1$ the limiting distribution of $\sqrt{T}(\hat{\underset{\sim}{B}} - \underset{\sim}{B})$ was proved in detail [Anderson (1971), Section 5.5, for example], and the method holds for any fixed N . Regardless of the distribution of $\underset{\sim}{y}_{1\alpha}$, if the $\underset{\sim}{u}_t$ are identically distributed $(t = 2,3,\ldots)$, $\sqrt{TN}(\hat{\underset{\sim}{B}} - \underset{\sim}{B})$ has a limiting normal distribution with mean $\underset{\sim}{0}$ and covariances $\underset{\sim}{\Sigma} \otimes \underset{\sim}{F}^{-1}$ as $T \to \infty$.

It will be noted that the limiting distribution of $\sqrt{N(T-1)}(\hat{\underset{\sim}{B}} - \underset{\sim}{B})$ in the stationary case is the same for T fixed and $N \to \infty$ and for N fixed and $T \to \infty$. We shall denote this situation as N and/or $T \to \infty$. Precisely this means that given an $\varepsilon > 0$ there exists an N_0 and a T_0 such that the normal distribution differs from the exact distribution by no more than ε if $N \geq N_0$ for any T or $T \geq T_0$ for any N . Of course, the difference is less than ε if both $N \geq N_0$ and $T \geq T_0$.

4. Tests of Hypotheses.

4.1. Introduction. There are various hypotheses about the autoregression matrices that one might want to test. Because the asymptotic distributions of the estimates of autoregression matrices are the same as the distributions of estimates of regression matrices when the dependent variables (analogous to $\underset{\sim}{y}_t$) are normally distributed and the independent variables (analogous to $\underset{\sim}{y}_{t-1}$) are nonstochastic, we can take over appropriate test procedures from normal regression theory [Chapter 8 of Anderson (1958)].

For many hypotheses there are choices of test criteria that have the same limiting distributions under the respective null hypotheses. In the normal regression with fixed independent variables a test procedure is based on a $p \times p$ matrix $\underset{\sim}{G}$ which has a Wishart distribution with covariance matrix $\underset{\sim}{\Sigma}$ and degrees of freedom n and a $p \times p$ matrix $\underset{\sim}{H}$ which has a noncentral Wishart distribution with covariance matrix $\underset{\sim}{\Sigma}$, degrees of freedom k, and a matrix-valued noncentrality parameter which is $\underset{\sim}{0}$ under the null hypothesis. The matrix $\underset{\sim}{G}$ is proportional to the maximum likelihood estimate $\hat{\underset{\sim}{\Sigma}}_{\Omega}$ of $\underset{\sim}{\Sigma}$ when the null hypothesis is true, and $\underset{\sim}{G} + \underset{\sim}{H}$ is proportional to the maximum likelihood estimate $\hat{\underset{\sim}{\Sigma}}_{\omega}$ of $\underset{\sim}{\Sigma}$ when the null hypothesis is not assumed true. Some criteria are a monotonic function of the Wilks likelihood ratio criterion $|\underset{\sim}{G}|/|\underset{\sim}{G} + \underset{\sim}{H}|$, the Lawley-Hotelling criterion $\operatorname{tr} \underset{\sim}{H}\underset{\sim}{G}^{-1}$, the Pillai criterion $\operatorname{tr} \underset{\sim}{H}(\underset{\sim}{G} + \underset{\sim}{H})^{-1}$, and the Roy maximum characteristic root of $\underset{\sim}{H}\,\underset{\sim}{G}^{-1}$ (or equivalently of $\underset{\sim}{H}(\underset{\sim}{G} + \underset{\sim}{H})^{-1}$). In this paper we

shall present $\operatorname{tr} \underset{\sim}{H}\, \underset{\sim}{G}^{-1}$ as the criterion because it seems most analogous to the corresponding criterion for Markov chains, but any of the others could be used. All the criteria have the same limiting χ^2-distribution when the null hypothesis is true.

4.2. Specified Autoregression Matrices. In the normal multivariate regression model treated in Chapter 8 of Anderson (1958) where the vector $\underset{\sim}{x}_\alpha$ is from $N(\underset{\sim}{\beta}\underset{\sim}{z}_\alpha, \underset{\sim}{\Sigma})$, $\alpha = 1, \ldots, N$, criteria for testing the null hypothesis $\underset{\sim}{\beta} = \underset{\sim}{\beta}^*$ depend on the matrices

$$\underset{\sim}{H} = (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}^*) \sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha \underset{\sim}{z}_\alpha' (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}^*)' \tag{4.1}$$

and

$$\begin{aligned} \underset{\sim}{G} = N \hat{\underset{\sim}{\Sigma}}_\Omega &= \sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha - \hat{\underset{\sim}{\beta}} \underset{\sim}{z}_\alpha)(\underset{\sim}{x}_\alpha - \hat{\underset{\sim}{\beta}} \underset{\sim}{z}_\alpha)' \\ &= \sum_{\alpha=1}^{N} \underset{\sim}{x}_\alpha \underset{\sim}{x}_\alpha' - \hat{\underset{\sim}{\beta}} \sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha \underset{\sim}{z}_\alpha' \hat{\underset{\sim}{\beta}} , \\ &= \sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha - \underset{\sim}{\beta}^* \underset{\sim}{z}_\alpha)(\underset{\sim}{x}_\alpha - \underset{\sim}{\beta}^* \underset{\sim}{z}_\alpha)' - (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}^*) \sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha \underset{\sim}{z}_\alpha' (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}^*)' , \end{aligned} \tag{4.2}$$

where

$$\hat{\underset{\sim}{\beta}} = \sum_{\alpha=1}^{N} \underset{\sim}{x}_\alpha \underset{\sim}{z}_\alpha' \left(\sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha \underset{\sim}{z}_\alpha' \right)^{-1} . \tag{4.3}$$

In testing the null hypothesis that $\underset{\sim}{B}(t) = \underset{\sim}{B}^{o}(t)$ for a given t $(t = 2, \ldots, T)$, where $\underset{\sim}{B}^{o}(t)$ is a completely specified matrix, the analogue of $\Sigma_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha} \underset{\sim}{z}_{\alpha}'$ is $\underset{\sim}{C}_{t-1}(0)$, the analogue of $\hat{\underset{\sim}{\beta}}$ is $\hat{\underset{\sim}{\beta}}(t)$, and the analogue of $\hat{\underset{\sim}{\Sigma}}_{\Omega}$ is $\hat{\underset{\sim}{\Sigma}}_{t}$. The trace criterion for testing the null hypothesis is

$$(4.4) \qquad N \operatorname{tr}[\hat{\underset{\sim}{B}}(t) - \underset{\sim}{B}^{o}(t)]\underset{\sim}{C}_{t-1}(0)[\hat{\underset{\sim}{B}}(t) - \underset{\sim}{B}^{o}(t)]'\hat{\underset{\sim}{\Sigma}}_{t}^{-1} .$$

Under the null hypothesis this criterion has a limiting χ^2-distribution with p^2 degrees of freedom as $N \to \infty$.

<u>Theorem</u>. If $\underset{\sim}{X}$ $(p \times p)$ has the distribution $N(\underset{\sim}{0}, \underset{\sim}{\Sigma} \otimes \underset{\sim}{F}^{-1})$, then $\operatorname{tr} \underset{\sim}{X} \underset{\sim}{F} \underset{\sim}{X}' \underset{\sim}{\Sigma}^{-1}$ has the χ^2-distribution with p^2 degrees of freedom.

<u>Proof</u>. Let $\underset{\sim}{F} = \underset{\sim}{K} \underset{\sim}{K}'$, where $\underset{\sim}{K}$ is $p \times p$ and nonsingular, and let $\underset{\sim}{V} = \underset{\sim}{X} K$. Then

$$(4.5) \qquad \operatorname{tr} \underset{\sim}{X} \underset{\sim}{F} \underset{\sim}{X}'\underset{\sim}{\Sigma}^{-1} = \operatorname{tr} \underset{\sim}{V} \underset{\sim}{V}'\underset{\sim}{\Sigma}^{-1} = \operatorname{tr} \underset{\sim}{V}'\underset{\sim}{\Sigma}^{-1} \underset{\sim}{V} ,$$

and $\underset{\sim}{V}$ is normally distributed with mean $\underset{\sim}{0}$; the covariance between the i-th and j-th rows of $\underset{\sim}{V}$ is $\sigma_{ij} \underset{\sim}{K}'\underset{\sim}{F}^{-1}\underset{\sim}{K} = \sigma_{ij} \underset{\sim}{I}$. The columns of $\underset{\sim}{V}$ are independently distributed, each according to $N(\underset{\sim}{0}, \underset{\sim}{\Sigma})$. Thus $\operatorname{tr} \underset{\sim}{V}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{V} = \Sigma_{j=1}^{p} \underset{\sim}{v}_{j} \underset{\sim}{\Sigma}^{-1} \underset{\sim}{v}_{j}$, where $\underset{\sim}{V} = (\underset{\sim}{v}_{1}, \ldots, \underset{\sim}{v}_{p})$, has a χ^2-distribution with p^2 degrees of freedom. Q.E.D.

Because the quantities $\sqrt{N}[\hat{\underset{\sim}{B}}(t) - \underset{\sim}{B}^{o}(t)]$ are independent in the limit for different t, the criteria (4.4) for different t are

independent. Hence, to test the null hypothesis $\underset{\sim}{B}(t) = \underset{\sim}{B}^{o}(t)$, $t = 2, \ldots, T$, the sum of (4.4) on t from $t = 2$ to $t = T$ can be used. Under the null hypothesis this has a limiting χ^2-distribution with $(T-2)p^2$ degrees of freedom.

If the autoregressive matrices are the same, that is, $\underset{\sim}{B}(1) = \ldots \underset{\sim}{B}(T) = \underset{\sim}{B}$, one can test the null hypothesis that $\underset{\sim}{B}$ is a specified $p \times p$ matrix $\underset{\sim}{B}^{o}$ by use of the criterion

$$N \operatorname{tr}(\hat{\underset{\sim}{B}} - \underset{\sim}{B}^{o}) \underset{\sim}{C}^{*}(0) (\hat{\underset{\sim}{B}} - \underset{\sim}{B}^{o})\hat{\underset{\sim}{\Sigma}}^{-1} . \tag{4.6}$$

As $N \to \infty$ and/or $T \to \infty$, this criterion has a limiting χ^2-distribution with p^2 degrees of freedom.

A set of test procedures can be inverted to obtain confidence regions; a confidence region for $\underset{\sim}{B}$ consists of all matrices $\underset{\sim}{B}^{o}$ such that (4.6) is less than a suitable number from the χ^2-tables.

<u>4.3. Equality of Autoregression Matrices given Equality of Covariance Matrices</u>. In case a process is not assumed stationary it may be of interest to ask whether the autoregression coefficients vary over time. We shall assume the distributions of the disturbances are the same, that is, $\underset{\sim}{\Sigma}_t = \underset{\sim}{\Sigma}$, $t = 2, \ldots, T$, for some covariance matrix $\underset{\sim}{\Sigma}$. Consider testing the null hypothesis $\underset{\sim}{B}(t) = \underset{\sim}{B}$, $t = 2, \ldots, T$, for some matrix $\underset{\sim}{B}$. The criterion

$$(4.7) \qquad N \operatorname{tr} \sum_{t=2}^{T} [\hat{\underset{\sim}{B}}(t) - \hat{\underset{\sim}{B}}] \underset{\sim}{C}_{t-1}(0) [\hat{\underset{\sim}{B}}(t) - \hat{\underset{\sim}{B}}]' \hat{\underset{\sim}{\Sigma}}^{-1}$$

$$= N \operatorname{tr} \left[\sum_{t=2}^{T} \hat{\underset{\sim}{B}}(t) \underset{\sim}{C}_{t-1}(0) \hat{\underset{\sim}{B}}(t)' - \hat{\underset{\sim}{B}} \underset{\sim}{C}^{*}(0) \hat{\underset{\sim}{B}}' \right] \hat{\underset{\sim}{\Sigma}}^{-1}$$

has a limiting χ^2-distribution with $(T-2)p^2$ degrees of freedom as $N \to \infty$ when the null hypothesis is true.

To justify the criterion, again consider the normal multivariate regression model where $\underset{\sim}{x}_{t\alpha}$ is from $N[\underset{\sim}{\beta}(t) \underset{\sim}{z}_{t\alpha}, \underset{\sim}{\Sigma}]$, $\alpha = 1, \ldots, N$, $t = 2, \ldots, T$. Then maximum likelihood estimates are

$$(4.8) \qquad \hat{\underset{\sim}{\beta}}(t) = \sum_{\alpha=1}^{N} \underset{\sim}{x}_{t\alpha} \underset{\sim}{z}'_{t\alpha} \left(\sum_{\alpha=1}^{N} \underset{\sim}{z}_{t\alpha} \underset{\sim}{z}'_{t\alpha} \right)^{-1}, \qquad t = 2, \ldots, T ;$$

$$(4.9) \qquad \hat{\underset{\sim}{\Sigma}} = \frac{1}{N(T-1)} \sum_{t=2}^{T} \sum_{\alpha=1}^{N} [\underset{\sim}{x}_{t\alpha} - \hat{\underset{\sim}{\beta}}(t) \underset{\sim}{z}_{t\alpha}][\underset{\sim}{x}_{t\alpha} - \hat{\underset{\sim}{\beta}}(t) \underset{\sim}{z}_{t\alpha}]' .$$

If $\underset{\sim}{\beta}(2) = \ldots = \underset{\sim}{\beta}(T) = \underset{\sim}{\beta}$, say, the maximum likelihood estimate of $\underset{\sim}{B}$ is

$$(4.10) \qquad \hat{\underset{\sim}{\beta}} = \sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{x}_{t\alpha} \underset{\sim}{z}'_{t\alpha} \left(\sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{z}_{t\alpha} \underset{\sim}{z}'_{t\alpha} \right)^{-1} .$$

In this case

$$(4.11) \qquad \underset{\sim}{H} = \sum_{t=2}^{T} [\hat{\underset{\sim}{\beta}}(t) - \underset{\sim}{\beta}] \sum_{\alpha=1}^{N} \underset{\sim}{z}_{t\alpha} \underset{\sim}{z}'_{t\alpha} [\hat{\underset{\sim}{\beta}}(t) - \underset{\sim}{\beta}]'$$

$$= \sum_{t=2}^{T} \hat{\underset{\sim}{\beta}}(t) \sum_{\alpha=1}^{N} \underset{\sim}{z}_{t\alpha} \underset{\sim}{z}'_{t\alpha} \hat{\underset{\sim}{\beta}}'(t) - \hat{\underset{\sim}{\beta}} \sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{z}_{t\alpha} \underset{\sim}{z}'_{t\alpha} \hat{\underset{\sim}{\beta}}'$$

has a Wishart distribution with covariance matrix $\underset{\sim}{\Sigma}$ and $(T-2)p^2$ degrees of freedom. The justification of this statement follows from Anderson (1958), Section 8.3, in a manner similar to Section 8.8. Then the trace of (4.11) times $\underset{\sim}{\Sigma}^{-1}$ has a χ^2-distribution with degrees of freedom equal to $T-1$ times the number of components in $\underset{\sim}{x}_{t\alpha}$ times the number of components in $\underset{\sim}{z}_{t\alpha}$. The criterion (4.7) is asymptotically similar.

4.4. Independence. The stochastic process $\{\underset{\sim}{y}_t\}$ consists of independent random vectors if $\underset{\sim}{B}(t) = \underset{\sim}{0}$ for all t in the inhomogeneous case and if $\underset{\sim}{B} = \underset{\sim}{0}$ in the homogeneous case. In the two cases independence corresponds to certain autoregression matrices being equal to $\underset{\sim}{0}$. Hence, appropriate criteria are the special cases of the criteria of Section 4.2 when $\underset{\sim}{B}^{o}(t) = \underset{\sim}{0}$ and $\underset{\sim}{B}^{o} = \underset{\sim}{0}$. To test the null hypothesis $\underset{\sim}{B}(t) = \underset{\sim}{0}$ the trace criterion is

$$N \operatorname{tr} \hat{\underset{\sim}{B}}(t)\, \underset{\sim}{C}_{t-1}(0)\, \hat{\underset{\sim}{B}}(t)'\, \hat{\Sigma}_t^{-1} . \tag{4.12}$$

To test $\underset{\sim}{B}(t) = 0$, $t = 2, \ldots, T$, one can use the sum of (4.12) for $t = 2, \ldots, T$ can be used. If $\underset{\sim}{\Sigma}_2 = \ldots = \underset{\sim}{\Sigma}_T$ is assumed $\hat{\underset{\sim}{\Sigma}}_t$ may be replaced by $\hat{\underset{\sim}{\Sigma}}$.

To test the null hypothesis of independence in a homogeneous process, that is, $\underset{\sim}{B} = \underset{\sim}{0}$ one can use the criterion

$$N \operatorname{tr} \hat{\underset{\sim}{B}}\, \underset{\sim}{C}^{*}(0)\, \hat{\underset{\sim}{B}}'\, \hat{\underset{\sim}{\Sigma}}^{-1} . \tag{4.13}$$

Note that the criterion for testing $\underset{\sim}{B}(2) = \ldots \underset{\sim}{B}(T) = \underset{\sim}{0}$ given $\underset{\sim}{\Sigma}_2 = \ldots = \underset{\sim}{\Sigma}_T$ is the sum of the criteria for testing $\underset{\sim}{B}(t) = \underset{\sim}{B}$, $t = 2, \ldots, T$, for some matrix $\underset{\sim}{B}$ and the criterion (4.13) for testing that $\underset{\sim}{B} = \underset{\sim}{0}$.

4.5. Test of a Given Order. The greater the order of an autoregressive process the further back in time is the dependence. In the interest of parsimony an investigator may ask whether the order of dependence is one integer, given that the order is not greater than another, which is larger. In Section 4.4 we considered a test of the hypothesis that the order of dependence was 0 given that the order was not greater than 1. In general, the hypothesis that a process is of a specified order given that it is of an order not greater than a certain integer is a hypothesis that the higher-order autoregression matrices are $\underset{\sim}{0}$.

To develop the procedures and (asymptotic) distribution theory we refer to the case of nonstochastic independent variables (Chapter 8 of Anderson (1958), for example). In the normal regression model $N(\underset{\sim}{\beta}\underset{\sim}{z}_\alpha, \underset{\sim}{\Sigma})$, $= 1, \ldots, N$, partition

$$\underset{\sim}{\beta} = (\underset{\sim}{\beta}_1 \; \underset{\sim}{\beta}_2) , \tag{4.14}$$

$$\underset{\sim}{z}_\alpha = \begin{pmatrix} \underset{\sim}{z}_\alpha^{(1)} \\ \underset{\sim}{z}_\alpha^{(2)} \end{pmatrix} . \tag{4.15}$$

Consider testing the null hypothesis $\omega : \underset{\sim}{\beta}_2 = \underset{\sim}{0}$. The criterion depends on the two matrices

$$(4.16) \qquad G = N\hat{\underset{\sim}{\Sigma}}_\Omega = \sum_{\alpha=1}^{N} (\underset{\sim}{x}_\alpha - \hat{\underset{\sim}{\beta}}_\Omega \underset{\sim}{z}_\alpha)(\underset{\sim}{x}_\alpha - \hat{\underset{\sim}{\beta}}_\Omega \underset{\sim}{z}_\alpha)'$$

$$= \sum_{\alpha=1}^{N} \underset{\sim}{x}_\alpha \underset{\sim}{x}_\alpha' - \hat{\underset{\sim}{\beta}}_\Omega \sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha \underset{\sim}{z}_\alpha' \hat{\underset{\sim}{\beta}}_\Omega' \quad ,$$

$$(4.17) \qquad N\hat{\underset{\sim}{\Sigma}}_\omega = \sum_{\alpha=1}^{N} \left(\underset{\sim}{x}_\alpha - \hat{\underset{\sim}{\beta}}_{1\omega} \underset{\sim}{z}_\alpha^{(1)}\right)\left(\underset{\sim}{x}_\alpha - \hat{\underset{\sim}{\beta}}_{1\omega} \underset{\sim}{z}_\alpha^{(1)}\right)'$$

$$= \sum_{\alpha=1}^{N} \underset{\sim}{x}_\alpha \underset{\sim}{x}_\alpha' - \hat{\underset{\sim}{\beta}}_{1\omega} \sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha^{(1)} \underset{\sim}{z}_\alpha^{(1)'} \hat{\underset{\sim}{\beta}}_{1\omega}' \ ,$$

where

$$(4.18) \qquad \hat{\underset{\sim}{\beta}}_\Omega = \sum_{\alpha=1}^{N} \underset{\sim}{x}_\alpha \underset{\sim}{z}_\alpha' \left(\sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha \underset{\sim}{z}_\alpha'\right)^{-1} \ ,$$

$$(4.19) \qquad \hat{\underset{\sim}{\beta}}_{1\omega} = \sum_{\alpha=1}^{N} \underset{\sim}{x}_\alpha \underset{\sim}{z}_\alpha^{(1)'} \left(\sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha^{(1)} \underset{\sim}{z}_\alpha^{(1)'}\right)^{-1} \ .$$

Then

$$
\begin{aligned}
(4.20) \qquad H &= N\hat{\underset{\sim}{\Sigma}}_{\omega} - N\hat{\underset{\sim}{\Sigma}}_{\Omega} \\
&= \hat{\underset{\sim}{\beta}}_{\Omega} \sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha} \underset{\sim}{z}_{\alpha}' \hat{\underset{\sim}{\beta}}_{\Omega} - \hat{\underset{\sim}{\beta}}_{1\omega} \sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha}^{(1)} \underset{\sim}{z}_{\alpha}^{(1)'} \hat{\underset{\sim}{\beta}}_{1\omega}' \\
&= \sum_{\alpha=1}^{N} \underset{\sim}{x}_{\alpha} \underset{\sim}{z}_{\alpha}' \left(\sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha} \underset{\sim}{z}_{\alpha}' \right)^{-1} \sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha} \underset{\sim}{x}_{\alpha}' \\
&\quad - \sum_{\alpha=1}^{N} \underset{\sim}{x}_{\alpha} \underset{\sim}{z}_{\alpha}^{(1)'} \left(\sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha}^{(1)} \underset{\sim}{z}_{\alpha}^{(1)'} \right)^{-1} \sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha}^{(1)} \underset{\sim}{x}_{\alpha}' . \\
&= \sum_{\alpha=1}^{N} \underset{\sim}{x}_{\alpha} \left[\underset{\sim}{z}_{\alpha}^{(2)} - \sum_{\beta=1}^{N} \underset{\sim}{z}_{\beta}^{(2)} \underset{\sim}{z}_{\beta}^{(1)'} \left(\sum_{\beta=1}^{N} \underset{\sim}{z}_{\beta}^{(1)} \underset{\sim}{z}_{\beta}^{(1)'} \right)^{-1} \underset{\sim}{z}_{\alpha}^{(1)} \right]' \\
&\quad \left\{ \sum_{\alpha=1}^{N} \left[\underset{\sim}{z}_{\alpha}^{(2)} - \sum_{\beta=1}^{N} \underset{\sim}{z}_{\beta}^{(2)} \underset{\sim}{z}_{\beta}^{(1)'} \left(\sum_{\beta=1}^{N} \underset{\sim}{z}_{\beta}^{(1)} \underset{\sim}{z}_{\beta}^{(1)'} \right)^{-1} \underset{\sim}{z}_{\alpha}^{(1)} \right] \right. \\
&\quad \left. \left[\underset{\sim}{z}_{\alpha}^{(2)} - \sum_{\beta=1}^{N} \underset{\sim}{z}_{\beta}^{(2)} \underset{\sim}{z}_{\beta}^{(1)'} \left(\sum_{\beta=1}^{N} \underset{\sim}{z}_{\beta}^{(1)} \underset{\sim}{z}_{\beta}^{(1)'} \right)^{-1} \underset{\sim}{z}_{\alpha}^{(1)} \right]' \right\}^{-1} \\
&\quad \sum_{\alpha=1}^{N} \left[\underset{\sim}{z}_{\alpha}^{(2)} - \sum_{\beta=1}^{N} \underset{\sim}{z}_{\beta}^{(2)} \underset{\sim}{z}_{\beta}^{(1)'} \left(\sum_{\beta=1}^{N} \underset{\sim}{z}_{\beta}^{(1)} \underset{\sim}{z}_{\beta}^{(1)'} \right)^{-1} \underset{\sim}{z}_{\alpha}^{(1)} \right] \underset{\sim}{x}_{\alpha}' \\
&= \hat{\beta}_{2\Omega} \left[\sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha}^{(2)} \underset{\sim}{z}_{\alpha}^{(2)'} - \sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha}^{(2)} \underset{\sim}{z}_{\alpha}^{(1)'} \left(\sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha}^{(1)} \underset{\sim}{z}_{\alpha}^{(1)'} \right)^{-1} \right. \\
&\qquad \left. \sum_{\alpha=1}^{N} \underset{\sim}{z}_{\alpha}^{(1)} \underset{\sim}{z}_{\alpha}^{(2)'} \right] \hat{\beta}_{2\Omega}' .
\end{aligned}
$$

Let us treat specifically the homogeneous case. We assume the order of dependence is r and we test the null hypothesis that the order is $q(<r)$; that is, the null hypothesis is

$$(4.21) \qquad \underset{\sim}{B}_{q+1} = \dots = \underset{\sim}{B}_r = \underset{\sim}{0} \ .$$

The regression matrices are estimated by (3.17) and $\underset{\sim}{\Sigma}$ is estimated by

$$(4.22) \qquad \hat{\Sigma} = \frac{1}{N(T-r)} \sum_{t=r+1}^{T} \sum_{\alpha=1}^{N} \left[\underset{\sim}{y}_{t\alpha} - \sum_{i=1}^{r} \hat{\underset{\sim}{B}}(i) \, \underset{\sim}{y}_{t-i,\alpha} \right] \left[\underset{\sim}{y}_{t\alpha} - \sum_{j=1}^{r} \hat{\underset{\sim}{B}}(j) \, \underset{\sim}{y}_{t-j,\alpha} \right]'$$

$$= C(0|r) - \sum_{i,j=1}^{r} \hat{\underset{\sim}{B}}(i) \, \underset{\sim}{C}(i,j|r) \, \hat{\underset{\sim}{B}}(j)' \ .$$

The criterion then is

(4.23)

$$N(T-r)\mathrm{tr} \sum_{i,j=q+1}^{r} \hat{\underset{\sim}{B}}(i) \left[\underset{\sim}{C}(i,j|r) - \sum_{g,h=1}^{q} \underset{\sim}{C}(i,g|r) \bar{\underset{\sim}{C}}(g,h|r) \underset{\sim}{C}(h,j|r) \right] \hat{\underset{\sim}{B}}(j)' \hat{\underset{\sim}{\Sigma}}^{-1} \ ,$$

where

$$(4.24) \qquad \begin{bmatrix} \bar{\underset{\sim}{C}}(1,1 \ r) & \dots & \bar{\underset{\sim}{C}}(1,g \ r) \\ \vdots & & \vdots \\ \bar{\underset{\sim}{C}}(g,1 \ r) & \dots & \bar{\underset{\sim}{C}}(g,g \ r) \end{bmatrix} = \begin{bmatrix} \underset{\sim}{C}(1,1 \ r) & \dots & \underset{\sim}{C}(1,g \ r) \\ \vdots & & \vdots \\ \underset{\sim}{C}(g,1 \ r) & \dots & \underset{\sim}{C}(g,g \ r) \end{bmatrix}^{-1} \ .$$

The criterion has a limiting χ^2-distribution with $(r-q)p^2$ degrees of freedom under the null hypothesis as N and/or $T \to \infty$.

4.6. Several Processes Identical. Several populations may be sampled. There may be different socio-economic statuses, different income levels, or individuals subject to different treatments. The investigator may ask whether the several processes are identical. Frequently it is reasonable to treat the error terms as distributed similarly in the different populations.

Consider testing the null hypothesis that the matrices of autoregressive coefficients of s first-order homogeneous autoregressive processes with identical covariance matrices $\underset{\sim}{\Sigma}$ are equal on the basis of N_h observed time series of length T from the h-th process, $h = 1, \ldots, s$. If $\hat{\underset{\sim}{B}}^{(h)}$ is the estimate of the matrix for the h-th process, $\hat{\underset{\sim}{B}}$ is the pooled estimate of the hypothetically equal matrices, and $\hat{\underset{\sim}{\Sigma}}$ is the pooled estimate of $\underset{\sim}{\Sigma}$, a criterion for testing the null hypothesis is

$$(4.25) \qquad \operatorname{tr} \sum_{h=1}^{s} (\hat{\underset{\sim}{B}}^{(h)} - \hat{\underset{\sim}{B}}) \sum_{t=2}^{T} \sum_{\alpha=1}^{N_h} \underset{\sim}{y}^{(h)}_{t-1,\alpha} \underset{\sim}{y}^{(h)'}_{t-1,\alpha} (\hat{\underset{\sim}{B}}^{(h)} - \hat{\underset{\sim}{B}})' \hat{\underset{\sim}{\Sigma}}^{-1}$$

$$= \operatorname{tr}\left[\sum_{h=1}^{s} \hat{\underset{\sim}{B}}^{(h)} \sum_{t=2}^{T} \sum_{\alpha=1}^{N_h} \underset{\sim}{y}^{(h)}_{t-1,\alpha} \underset{\sim}{y}^{(h)}_{t-1,\alpha} \hat{\underset{\sim}{B}}^{(h)'} - \hat{\underset{\sim}{B}} \sum_{h=1}^{s} \sum_{t=2}^{T} \sum_{\alpha=1}^{N_h} \underset{\sim}{y}^{(h)}_{t-1,\alpha} \underset{\sim}{y}^{(h)}_{t-1,\alpha} \hat{\underset{\sim}{B}}'\right] \hat{\underset{\sim}{\Sigma}}^{-1} .$$

Under the null hypothesis this has a limiting χ^2-distribution with

$(s-1)p^2$ degrees of freedom as N and/or $T \to \infty$. The justification of the procedure is similar to that of Section 4.3.

In the inhomogeneous case we can test the null hypothesis $\underset{\sim}{B}^{(1)}(t) = \ldots = \underset{\sim}{B}^{(2)}(t)$ for a given t by use of (4.25) with $\hat{\underset{\sim}{B}}^{(h)}$ replaced by $\hat{\underset{\sim}{B}}^{(h)}(t)$ and $\hat{\underset{\sim}{B}}$ replaced by $\hat{\underset{\sim}{B}}(t)$; the sums on t are deleted. To test the hypothesis for all $t = 2, \ldots, T$, the criteria are summed over t.

<u>4.6. Independence of Two Subprocesses</u>. Suppose $\underset{\sim}{y}_t = (\underset{\sim}{y}_t^{(1)'} \underset{\sim}{y}_t^{(2)'})'$, where $\underset{\sim}{y}_t^{(1)}$ and $\underset{\sim}{y}_t^{(2)}$ consist of p_1 and $p_2 (p_1 + p_2 = p)$ components, respectively. We ask whether the two subprocesses $\{\underset{\sim}{y}_t^{(1)}\}$ and $\{\underset{\sim}{y}_t^{(2)}\}$ are independent. Since the joint distributions of all random variables are normal with means 0, the subprocesses are independent if and only if

$$\mathcal{E}_{\underset{\sim}{y}_t^{(1)}} \underset{\sim}{y}_s^{(2)'} = \underset{\sim}{0} \tag{4.26}$$

for all t and s. Let

$$\underset{\sim}{B} = \begin{pmatrix} \underset{\sim}{B}_{11} & \underset{\sim}{B}_{12} \\ \underset{\sim}{B}_{21} & \underset{\sim}{B}_{22} \end{pmatrix}, \quad \underset{\sim}{\Sigma} = \begin{pmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{\Sigma}_{12} \\ \underset{\sim}{\Sigma}_{21} & \underset{\sim}{\Sigma}_{22} \end{pmatrix}. \tag{4.27}$$

The two subprocesses (in the Gaussian case) are independent if

$$\underset{\sim}{B}_{12} = \underset{\sim}{0}, \ \underset{\sim}{B}_{21} = \underset{\sim}{0}, \tag{4.28}$$

(4.29) $$\underset{\sim}{\Sigma}_{12} = \underset{\sim}{\Sigma}_{21}' = \underset{\sim}{0} ,$$

because then

(4.30)
$$\underset{\sim}{F} = \mathcal{E}\begin{pmatrix} \underset{\sim}{y}_t^{(1)} \\ \underset{\sim}{y}_t^{(2)} \end{pmatrix} (\underset{\sim}{y}_t^{(1)\prime} \underset{\sim}{y}_t^{(2)\prime}) = \sum_{s=0}^{\infty} \begin{pmatrix} \underset{\sim}{B}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{B}_{22} \end{pmatrix}^s \begin{pmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{\Sigma}_{22} \end{pmatrix} \begin{pmatrix} \underset{\sim}{B}_{11}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{B}_{22} \end{pmatrix}^s$$

$$= \begin{pmatrix} \sum_{s=0}^{\infty} B_{11}^s \, \Sigma_{11} \, B_{11}'^s & \underset{\sim}{0} \\ \underset{\sim}{0} & \sum_{s=0}^{\infty} \underset{\sim}{B}_{22}^s \, \underset{\sim}{\Sigma}_{22} \, \underset{\sim}{B}_{22}'^s \end{pmatrix}$$

$$= \begin{pmatrix} \underset{\sim}{F}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{F}_{22} \end{pmatrix} ,$$

and for $t \geq s$

(4.31) $$\mathcal{E}\begin{pmatrix} \underset{\sim}{y}_t^{(1)} \\ \underset{\sim}{y}_t^{(2)} \end{pmatrix} (\underset{\sim}{y}_s^{(1)\prime} \underset{\sim}{y}_s^{(2)\prime})$$

$$= \begin{pmatrix} \underset{\sim}{B}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{B}_{12} \end{pmatrix}^{t-s} \begin{pmatrix} \underset{\sim}{F}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{F}_{22} \end{pmatrix}$$

$$= \begin{pmatrix} \underset{\sim}{B}_{11}^{t-s} \, \underset{\sim}{F}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{B}_{22}^{t-s} \, \underset{\sim}{F}_{22} \end{pmatrix} .$$

On the other hand if $\underset{\sim}{y}_t^{(1)}$ and $\underset{\sim}{y}_t^{(2)}$ are to be independent $\underset{\sim}{F}_{12} = \underset{\sim}{F}_{21}' = \underset{\sim}{0}$ and if $\underset{\sim}{y}_{t+1}^{(1)}$ and $\underset{\sim}{y}_t^{(2)}$ are to be independent and

$\underset{\sim}{y}_{t+1}^{(2)}$ and $\underset{\sim}{y}_t^{(1)}$ are to be independent, then

$$(4.32)\qquad \mathcal{E}\begin{pmatrix} \underset{\sim}{y}_{t+1}^{(1)} \\ \underset{\sim}{y}_{t+1}^{(2)} \end{pmatrix} (\underset{\sim}{y}_t^{(1)\prime}\ \underset{\sim}{y}_t^{(2)\prime}) = \begin{pmatrix} \underset{\sim}{B}_{11} & \underset{\sim}{B}_{12} \\ \underset{\sim}{B}_{21} & \underset{\sim}{B}_{22} \end{pmatrix} \begin{pmatrix} \underset{\sim}{F}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{F}_{22} \end{pmatrix}$$

$$= \begin{pmatrix} \underset{\sim}{B}_{11}\underset{\sim}{F}_{11} & \underset{\sim}{B}_{12}\underset{\sim}{F}_{22} \\ \underset{\sim}{B}_{21}\underset{\sim}{F}_{11} & \underset{\sim}{B}_{22}\underset{\sim}{F}_{22} \end{pmatrix}$$

$$= \begin{pmatrix} \underset{\sim}{B}_{11}\underset{\sim}{F}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{B}_{22}\underset{\sim}{F}_{22} \end{pmatrix}.$$

Hence, (4.28) must hold. Then

$$(4.33)\qquad \underset{\sim}{F} = \begin{pmatrix} \underset{\sim}{F}_{11} & \underset{\sim}{F}_{12} \\ \underset{\sim}{F}_{21} & \underset{\sim}{F}_{22} \end{pmatrix} = \sum_{s=0}^{\infty} \begin{pmatrix} \underset{\sim}{B}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{B}_{22} \end{pmatrix}^s \begin{pmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{\Sigma}_{12} \\ \underset{\sim}{\Sigma}_{21} & \underset{\sim}{\Sigma}_{22} \end{pmatrix} \begin{pmatrix} \underset{\sim}{B}_{11}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{B}_{22}' \end{pmatrix}^s$$

$$= \begin{pmatrix} \sum_{s=0}^{\infty} \underset{\sim}{B}_{11}^s \underset{\sim}{\Sigma}_{11} \underset{\sim}{B}_{11}'^s & \sum_{s=0}^{\infty} \underset{\sim}{B}_{11}^s \underset{\sim}{\Sigma}_{12} \underset{\sim}{B}_{22}'^s \\ \sum_{s=0}^{\infty} \underset{\sim}{B}_{22}^s \underset{\sim}{\Sigma}_{21} \underset{\sim}{B}_{11}'^s & \sum_{s=0}^{\infty} \underset{\sim}{B}_{22}^s \underset{\sim}{\Sigma}_{22} \underset{\sim}{B}_{22}'^s \end{pmatrix}$$

$$= \begin{pmatrix} \underset{\sim}{F}_{11} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{F}_{22} \end{pmatrix}$$

implies (4.29). (See Section 5.3 of Anderson (1971).)

To test the hypothesis (4.29) we use

$$\hat{\underset{\sim}{\Sigma}} = \begin{pmatrix} \hat{\underset{\sim}{\Sigma}}_{11} & \hat{\underset{\sim}{\Sigma}}_{12} \\ \hat{\underset{\sim}{\Sigma}}_{21} & \hat{\underset{\sim}{\Sigma}}_{22} \end{pmatrix} . \tag{4.34}$$

We develop the test criterion in analogy to the usual normal regression case. If $\underset{\sim}{x}_\alpha = (\underset{\sim}{x}_\alpha^{(1)'} \underset{\sim}{x}_\alpha^{(2)'})'$ has the distribution $N(\underset{\sim}{0}, \underset{\sim}{\Sigma})$, $\alpha = 1, .., n$, where Σ is partitioned according to (4.27), then the conditional distribution of $\underset{\sim}{x}_\alpha^{(1)}$ given $\underset{\sim}{x}_\alpha^{(2)}$ is $N(\underset{\sim}{\beta} z_\alpha^{(2)}, \Sigma_{11.2})$, where $\underset{\sim}{\beta} = \underset{\sim}{\Sigma}_{12} \underset{\sim}{\Sigma}_{22}^{-1}$ and $\underset{\sim}{\Sigma}_{11.2} = \underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Sigma}_{11} - \underset{\sim}{\Sigma}_{12} \underset{\sim}{\Sigma}_{22}^{-1} \underset{\sim}{\Sigma}_{21}$. If

$$\underset{\sim}{A} = \sum_{\alpha=1}^{n} \underset{\sim}{x}_\alpha \underset{\sim}{x}_\alpha' = \begin{pmatrix} \underset{\sim}{A}_{11} & \underset{\sim}{A}_{12} \\ \underset{\sim}{A}_{21} & \underset{\sim}{A}_{22} \end{pmatrix} , \tag{4.35}$$

then the estimate of $\underset{\sim}{\beta}$ is $\underset{\sim}{A}_{12} \underset{\sim}{A}_{22}^{-1}$ and the estimate of $\underset{\sim}{\Sigma}_{11.2}$ is $(1/n)(\underset{\sim}{A}_{11} - \underset{\sim}{A}_{12} \underset{\sim}{A}_{22}^{-1} \underset{\sim}{A}_{21})$. The trace test for $\underset{\sim}{\beta} = \underset{\sim}{0}$ is

$$n \operatorname{tr} \underset{\sim}{A}_{12} \underset{\sim}{A}_{22}^{-1} \underset{\sim}{A}_{21} (\underset{\sim}{A}_{11} - \underset{\sim}{A}_{12} \underset{\sim}{A}_{22}^{-1} \underset{\sim}{A}_{21})^{-1} ; \tag{4.36}$$

since $(1/n) \underset{\sim}{A}_{12}$ is a consistent estimate of $\underset{\sim}{\Sigma}_{12} = \underset{\sim}{0}$ when the null hypothesis is true, the criterion can be simplified to

$$n \operatorname{tr} \underset{\sim}{A}_{12} \underset{\sim}{A}_{22}^{-1} \underset{\sim}{A}_{21} \underset{\sim}{A}_{11}^{-1} . \tag{4.37}$$

To return to testing the hypothesis (4.29) for the autoregressive process, we can use the criterion

$$(4.38) \qquad N(T-1)\ \mathrm{tr}\ \hat{\Sigma}_{12}\ \hat{\Sigma}_{22}^{-1}\ \hat{\Sigma}_{21}\ \hat{\Sigma}_{11}^{-1}\ ,$$

which has a limiting χ^2-distribution with $p_1 p_2$ degrees of freedom as N and/or $T \to \infty$.

Now consider testing the null hypothesis $\underset{\sim}{B}_{12} = \underset{\sim}{0}$ when $\underset{\sim}{\Sigma}_{12} = \underset{\sim}{\Sigma}_{21}' = \underset{\sim}{0}$. As a guide we look again to the usual regression model $N(\underset{\sim}{\beta}\underset{\sim}{z}_\alpha, \underset{\sim}{\Sigma})$, where $\underset{\sim}{x}_\alpha = (\underset{\sim}{x}_\alpha^{(1)'}, \underset{\sim}{x}_\alpha^{(2)'})'$, $\underset{\sim}{z}_\alpha' = (\underset{\sim}{z}_\alpha^{(1)'}, \underset{\sim}{z}_\alpha^{(2)'})'$,

$$(4.39) \qquad \underset{\sim}{\beta} = \begin{pmatrix} \underset{\sim}{\beta}_{11} & \underset{\sim}{\beta}_{12} \\ \underset{\sim}{\beta}_{21} & \underset{\sim}{\beta}_{22} \end{pmatrix}, \quad \underset{\sim}{\Sigma} = \begin{pmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{\Sigma}_{12} \\ \underset{\sim}{\Sigma}_{21} & \underset{\sim}{\Sigma}_{22} \end{pmatrix}$$

and $\underset{\sim}{\Sigma}_{12} = \underset{\sim}{\Sigma}_{21}' = \underset{\sim}{0}$. Then the estimate of $(\underset{\sim}{\beta}_{11}\ \underset{\sim}{\beta}_{12})$ is distributed independently of $(\underset{\sim}{\beta}_{21}\ \underset{\sim}{\beta}_{22})$. In particular, $\hat{\underset{\sim}{\beta}}_{12}$ is normally distributed with mean $\underset{\sim}{\beta}_{12}$ and covariances constituting $\underset{\sim}{\Sigma}_{11} \otimes \underset{\sim}{D}_{22.1}^{-1}$, where

$$(4.40) \qquad \underset{\sim}{D}_{22.1} = \sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha^{(2)} \underset{\sim}{z}_\alpha^{(2)'} - \sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha^{(2)} \underset{\sim}{z}_\alpha^{(1)'} \left(\sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha^{(1)} \underset{\sim}{z}_\alpha^{(1)'} \right)^{-1} \sum_{\alpha=1}^{N} \underset{\sim}{z}_\alpha^{(1)} \underset{\sim}{z}_\alpha^{(2)'} .$$

The test criterion is an appropriate multiple of $\mathrm{tr}\ \hat{\underset{\sim}{\beta}}_{12}\ \underset{\sim}{D}_{22.1}\ \hat{\underset{\sim}{\beta}}_{21}\ \hat{\underset{\sim}{\Sigma}}_{11}^{-1}$.

For the autoregressive process we use

$$(4.41) \qquad \mathrm{tr}\ \hat{\underset{\sim}{B}}_{12} \left[\sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{y}_{t-1,\alpha}^{(2)} \underset{\sim}{y}_{t-1,\alpha}^{(2)'} - \sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{y}_{t-1,\alpha}^{(2)} \underset{\sim}{y}_{t-1,\alpha}^{(1)'} \left(\sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{y}_{t-1,\alpha}^{(1)} \underset{\sim}{y}_{t-1,\alpha}^{(1)'} \right)^{-1} \sum_{t=2}^{T} \sum_{\alpha=1}^{N} \underset{\sim}{y}_{t-1,\alpha}^{(1)} \underset{\sim}{y}_{t-1,\alpha}^{(2)'} \right] \hat{\underset{\sim}{B}}_{12}' \hat{\underset{\sim}{\Sigma}}_{11}^{-1} ;$$

its limiting distribution as N and/or $T \to \infty$ is a χ^2-distribution with $p_1 p_2$ degrees of freedom. A criterion to test the null hypothesis $\underset{\sim}{B}_{21} = \underset{\sim}{0}$ is (4.41) with 1 and 2 interchanged; its limiting distribution is also a χ^2-distribution with $p_1 p_2$ degrees of freedom. The three criteria are asymptotically independent.

5. Multiple Decision Problems.

5.1. General Comments. In a statistical investigation it frequently occurs that several problems are investigated simultaneously on the basis of one set of data. The designation of distinct hypotheses to test is a convenient division of a statistical study. There are some other aspects which are more conveniently viewed as making a choice or decision among several alternatives.

5.2. Selection of the Order of an Autoregressive Process. Suppose the maximum assumed order of an autoregressive homogeneous process is r and the minimum is q $(<r)$. It is desired to determine whether the order is $q, q+1, \ldots, r$. A possible procedure is to test whether the order is $r-1$ against the alternative that it is r at significance level ε_r. If this null hypothesis is accepted, then test the next null hypothesis that the order is $r-1$ against the alternative that it is $r-2$ at significance level ε_{r-1}, etc. If the orders $r-1, r-2, \ldots, q$ are accepted in sequence, then the order q is determined. For more discussion of this type of procedure see Anderson (1971), Section 6.4.

5.3. Equality of Regression Matrices over Intervals of Time. In the case of an inhomogeneous autoregressive process it is possible that the autoregression matrices are not equal over the entire range of time $t = 2, \ldots, T$, but are equal over intervals of time while being unequal between different intervals. The intervals may be preassigned. The method of Section 4.3 can be used to test simultaneously

the hypotheses

$$(5.1) \qquad \underset{\sim}{B}(2) = \ldots = \underset{\sim}{B}(t_1),\ \underset{\sim}{B}(t_1+1) = \ldots = \underset{\sim}{B}(t_2),\ \ldots,$$

$$\underset{\sim}{B}(t_{q-1}+1) = \ldots = \underset{\sim}{B}(t_q)\ .$$

The q criteria are asymptotically independent.

In T. W. Anderson (1954) are examples of the statistical analysis of panel data where the responses were discrete and the Markov chain served as a model. The transition probability matrix expressed the probability of a political party preference at one interview given the preference at the preceding (monthly) interview. One matrix fitted the data for May to July (before the Democratic convention), one for July to August (between the two conventions), and one for August to October (after the Republican convention).

If the autoregression matrices are not constant for $t = 2, \ldots, T$, there may be other patterns to them. If the characteristic roots of $\underset{\sim}{B}$ are different (or if the elementary divisions are simple), then $\underset{\sim}{B} = \underset{\sim}{Q}\underset{\sim}{\Lambda}\underset{\sim}{Q}^{-1}$, where $\underset{\sim}{\Lambda}$ is the diagonal matrix of characteristic roots and $\underset{\sim}{Q}$ is a matrix of characteristic vectors. In a homogeneous autoregressive process the conditional expectation of $\underset{\sim}{y}_t$ given $\underset{\sim}{y}_{t-2}$ is

$$(5.2) \qquad \mathcal{E}y_t|\underset{\sim}{y}_{t-2} = \underset{\sim}{B}^2\, y_{t-2} = \underset{\sim}{Q}\,\Lambda^2\,\underset{\sim}{Q}^{-1}\underset{\sim}{y}_t\ .$$

If a process is not homogeneous, we can ask whether $\underset{\sim}{B}(t) = \underset{\sim}{Q}\,\underset{\sim}{\Lambda}^{a(t)}\underset{\sim}{Q}^{-1}$ for some $\underset{\sim}{Q}$, diagonal $\underset{\sim}{\Lambda}$, and suitable numbers $a(2), \ldots, a(T)$. The sequence $1, a(2), a(2) + a(3), \ldots,$ could denote an "inherent" time.

6. Further Discussion. The model for the Kalman filter is usually a generalization of the inhomogeneous process considered here. The vector $\underset{\sim}{y}_t$ is an unobservable state vector. The observed measurement vector is

$$\underset{\sim}{z}_t = \underset{\sim}{H}(t)\, \underset{\sim}{y}_t + \underset{\sim}{v}_t \quad , \tag{6.1}$$

where $\underset{\sim}{v}_t$ is a random vector. In engineering applications the matrices $\underset{\sim}{B}(t)$, $\underset{\sim}{\Sigma}_t$, $\underset{\sim}{H}(t)$ and the covariance matrix of $\underset{\sim}{v}_t$ are usually known.

REFERENCES

Anderson, T. W. (1954), "Probability Models for Analyzing Time Changes in Attitudes", Mathematical Thinking in the Social Sciences, Paul F. Lazarsfeld, ed., The Free Press, Glencoe, Illinois, pp. 17-66.

Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, Inc., New York.

Anderson, T. W. (1971), The Statistical Analysis of Time Series, John Wiley and Sons, Inc., New York.

Anderson, T. W. (1976), "Panels and Time Series Analysis: Markov Chains and Autoregressive Processes", Technical Report No. 24, ONR Contract N00014-75-C-0442, Department of Statistics, Stanford University.

Anderson, T. W., and Goodman, Leo A. (1957), "Statistical Inference about Markov Chains", Annals of Mathematical Statistics, Vol. 28, pp. 89-110.

Appendix

This technical report gives the mathematical justification for assertions in Technical Report No. 24, elaborates on some aspects, and presents some new methods. This opportunity is taken to make some corrections and add an Addendum.

Corrections to Technical Report No. 24. Panels and Time Series Analysis: Markov Chains and Autoregressive Processes.

page 11, line 1	Replace $[p_{ij}(t)$ by $p_{ij}(t)[$.
page 16, line 1 under (3.7)	Replace $\underset{\sim}{B}(1)$ by $\underset{\sim}{B}(2)$.
page 17, line 6	Add $\Sigma_{t=2}^{T} \Sigma_{\alpha=1}^{N} \underset{\sim}{y}_{t-1,\alpha} \underset{\sim}{y}'_{t-1,\alpha}$.
page 22, (4.6)	Replace p_{ijk} by $\hat{p}_{ijk}$.
page 25, line 7	Replace $\underset{\sim}{B}(1)$ by $\underset{\sim}{B}(2)$.
page 25, (4.15), line 2	Replace $\underset{\sim}{B}$ by $\hat{\underset{\sim}{B}}$ 4 times.
page 26, line 4	Delete "given $\underset{\sim}{\Sigma}_2 = \ldots = \underset{\sim}{\Sigma}_T$" .
page 26, line 6	After "criterion" add "(given $\underset{\sim}{\Sigma}_2 = \ldots = \underset{\sim}{\Sigma}_T$)".
page 27, line 2 from bottom	Insert "limiting" before χ^2 .

Addendum: Technical Report No. 24 was written for a volume to celebrate the Seventy-fifth Birthday of Paul F. Lazarsfeld. Tragically, Professor Lazarsfeld died before the organization and publication of this volume was completed. The current technical report as well as the preceding are dedicated to his memory.

TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.

2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.

3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.

4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.

5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.

6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.

7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.

8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.

9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.

10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.

11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.

12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.

13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.

14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.

15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.

16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time Series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations," Steffen L. Lauritzen, May, 1974.

18. "On the Interrelationships among Sufficiency, Total Sufficiency and Some Related Concepts," Steffen L. Lauritzen, June, 1974.

19. "Statistical Inference for Multiply Truncated Power Series Distributions," T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models in the Time and Frequency Domains," T. W. Anderson, June 1975.

21. "Asymptotic Properties of Some Estimators in Moving Average Models," Raul Pedro Mentz, September 8, 1975.

22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting," Mituaki Huzii, February 1976.

23. "Estimating Means when Some Observations are Classified by Linear Discriminant Function," Chien-Pai Han, April 1976.

24. "Panels and Time Series Analysis: Markov Chains and Autoregressive Processes," T. W. Anderson, July 1976.

25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson, September 1976.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 25	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) REPEATED MEASUREMENTS ON AUTOREGRESSIVE PROCESSES.		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Theodore W. Anderson		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0442
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (NR-042-034)
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Statistics & Probability Program Code 436 Arlington, Virginia 22217		12. REPORT DATE September 1976
		13. NUMBER OF PAGES 40
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Autoregressive processes, time series analysis, repeated measurements, chi-square tests.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Estimation of parameters and tests of hypotheses are studied in first-order autoregressive processes where the process is observed several times over a given time interval. The process may be homogeneous (that is, the parameters may be constant over time) or inhomogeneous (time-varying parameters). Sufficient statistics under normality are obtained for various cases and several tests of hypotheses are given.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)